DOCUMENT RESUME

ED 061 636 24 EA 004 241

AUTHOR	Weaver, W. Timothy
TITLE	Delphi, A Critical Review. A Research Report.
INSTITUTION	Syracuse Univ. Research Corp., N.Y. Educational Policy Research Center.
SPONS AGENCY	National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
REPORT NO	EPRC-RR-7
BUREAU NO	BR-7-0996
PUB DATE	Feb 72
CONTRACT	OEC-1-7-070996-4253
NOTE	67p.
AVAILABLE FROM	Syracuse University Research Corporation, Merrill Lane, University Heights, Syracuse, New York 13210
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Decision Making Skills; Educational Policy; Educational Research; Feedback; Group Dynamics; Literature Reviews; Policy Formation; *Prediction; Public Policy; *Questioning Techniques; *Research Methodology; *Research Tools; Teaching Techniques
IDENTIFIERS	*Delphi Technique

ABSTRACT

The Delphi technique is a questionnaire method for organizing and shaping opinion through feedback. Although Delphi was originally intended as a tool for scientific and technological forecasting, its more promising educational application appears to be as (1) a method for studying the process of thinking about the future, (2) a pedagogical or teaching tool that forces people to think about the future in a more complex way than they would ordinarily, and (3) a planning tool that could aid in probing priorities held by members and constituencies of an organization. Related documents are EA 004 239 and EA 004 240. (Author/RA)

ED 061636

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

EPRC Research Report RR-7

DELPHI, A CRITICAL REVIEW

by

W. Timothy Weaver

Educational Policy Research Center
Syracuse University Research Corporation
1206 Harrison Street
Syracuse, New York 13210

EA 004 241

February 1972

Documents are available from the Educational Policy Research Center at Syracuse in three formats, besides the regular publication, Notes on the Future of Education:

RESEARCH REPORTS

Reports which have completed review by the EPRC and which deal with specific, policy oriented research. The reports in this series are usually marked by intensive research, either quantified or historical, and address themselves to specific research questions.

EXPLORATORY REPORTS

Reports which, while dealing with policy issues, often approach the realm of conjecture; they address themselves to social issues and the future, may be prescriptive rather than descriptive in tone, and are, by nature, more controversial in their conclusions. The review of these reports by the EPRC is as rigorous as that for Research Reports, though the conclusions remain those of the researcher rather than necessarily representing consensus agreement among the entire Center staff.

WORKING DRAFTS

Working Drafts are papers in progress, and are occasionally made available, in limited supply, to portions of the public to allow critical feedback and review. They have gone through little or no organized review at the Center, and their substance could reflect either of the above two categories of reports.

The research for this paper was conducted pursuant to Contract No. OEC-1-7-070996-4253 with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

TABLE OF CONTENTS

	Page
Preface	i
I. What Delphi Is	1
The Exploratory Delphi	1
Normative Forecasting	3
Family of Forecasting Tools	5
Explanatory Power	7
II. Applications of Delphi in Education	9
Normative Studies	9
Exploratory Studies	13
Some Criticisms	19
III. Examining the Research Background on Delphi	22
Future Cognition: Two Traditions of Research	22
The Early Studies: Forecasting and Forecasters	24
More Recent Studies on Forecasters and Forecasting	28
Feedback Studies	31
Personality Influence	34
EPRC Research	35
Notes on the Recent Research	36
Some Conclusions	41
IV. Modifying Delphi	43
Rationale	43
Notes	50
Delphi Studies in Education: Selected References	58
Selected Related Studies	60

PREFACE

Delphi, like the future it was intended to foretell, has not turned out to be what we expected. It displays certain fundamental weaknesses in its present form as a forecasting tool. Briefly, they have to do with interpreting the significance of convergence of opinion under conditions imposed by Delphi. The observation that people tend to shift their estimates toward a group norm under conditions of iteration is, on the basis of several controlled experiments with Delphi, a consistent and solid observation. There is some very meager evidence which suggests that compression of estimates over rounds produces a final consensus closer to the "true" answer (when the consensus is taken as a median of the spread of estimates). This finding, however, is based upon evidence collected from very short-term predictions in the economic domain, and from experiments with almanac-type questions. Just how accurately the findings can be generalized to Delphis which cover a 30-year extension into the future is unknown. Moreover, to make such a generalization is irrelevant to an understanding of plausibility in forecasting. Yet, interpreting the social-psychological significance of the convergency that does occur is important in understanding how the mind processes information about the future. Once we can understand more clearly how the mind formulates images of the future, we will be in a better position to improve upon the process of constructing rational and plausible forecasts.

At present Delphi forecasts come up short because there is little emphasis on the grounds or arguments which might convince policy-makers of the forecasts' reasonableness. There are insufficient procedures to

distinguish hope from likelihood. Delphi at present can render no rigorous distinction between reasonable judgment and mere guessing; nor does it clearly distinguish priority and value statements from rational arguments, nor feelings of confidence and desirability from statements of probability.

Of equally great importance, however, our research also leads us to conclude that Delphi, in combination with other tools, is a very potent device for teaching people to think about the future in much more complex ways than they ordinarily would. When we understand this use of Delphi we may find that, as a general teaching strategy, it is useful and more important than as a forecasting device. What this means is that initially the way we want to get educators (in our case) to make better decisions--decisions which account for alternative consequences--is to help them think in more complex ways about the future. Delphi seems ideally suited to such a purpose. Indeed, educators may find in Delphi and other forecasting tools a better pedagogy. One should not assume, however, that the weaknesses inherent in Delphi as a forecasting tool are its redeeming features as a teaching tool. Those weaknesses must be corrected if Delphi is to have any use at all, heuristic or otherwise.

Although Delphi was originally intended as a tool for scientific and technological forecasting, its more promising educational application seems to be in the following areas: (a) a method for studying the process of thinking about the future, (b) a pedagogical tool or teaching tool which forces people to think about the future in a more complex way than they ordinarily would, and (c) a planning tool which may aid in probing priorities held by members and constituencies of an organization.

DELPHI, A CRITICAL REVIEW

I.

WHAT DELPHI IS

The presumption that one mark of the creative man is an ingenious ability to play variations on a theme has never been more pronounced than in Delphi studies. Hundreds of interludes have followed the Rand Corporation's original composition. Although one never conducts a Delphi (one always conducts a "modified" Delphi), certain basic concepts have been preserved. In review these follow.

The Exploratory Delphi

The Delphi technique is a questionnaire method for organizing and shaping opinion through feedback. Its original use was to question experts as to their views about a chronology of scientific and technological events, and particularly to collect their judgments as to just when the events might occur.[1] Delphi has been justified primarily on the grounds that it prevents professional status and high position from forcing judgments in certain directions--as frequently occurs when panels of experts meet. The intention was to assure that changes in estimates reflected rational judgment, not the influence of certain opinion leaders. We will return to this point later.

Typically, the procedure includes a questionnaire mailed to respondents

who remain anonymous to one another. Respondents first generate several rather concise statements of events (or in some cases start out with the events already stated for them), and in the second round, give estimates as the probability of each event occurring at a given date in the future. Once the respondents have given their answers, the responses are collated and returned to each respondent who then is invited to revise his estimates. The third round responses are made with the knowledge of how others felt regarding the occurrence of each event. Again, the responses are assembled and reported back to the participants. If a respondent's estimate does not fall within the interquartile range of all conjectures, he is asked to justify his position, whether or not he wishes to change his position.

More recently, the technique has been extended to include questions about how familiar the participants are with the events. Respondents are also occasionally asked to rate the desirability of the events, should they occur. In addition, respondents are asked to give some statements about what impacts the events might have, if they occur. Still another question now being asked is what possible "interventions" might be developed to either enhance or reduce the probability that an event would occur.[2]

A number of variations have been played on this theme, but essentially they all end up asking a panel (sometimes referred to as experts, sometimes not) to assign dates or probabilities or both to rather specifically stated future events. In one way or another, the dates and probabilities of other members are revealed. The form of that revelation is usually such that a majority opinion is conveyed--taking for example, the median and inter-quartiles, or the average of the group, or the mode of the distribution of responses, as the majority opinion.

Regardless of the form and means used to establish the opinion feedback, the purpose of Delphi is to engage people in conjecturing about the <u>likelihood</u> of an event occurring at a particular time in the future. It is deliberately intended in these studies that the nature of that conjecture

2

be shaped and changed by the feedback of opinions of others until a point of relative stability is reached.

Normative Forecasting

The basic idea in the exploratory Delphi, the deliberate shaping of judgments through informative feedback, has been uprooted and transplanted in experiments with goal formulation.[3] This use of Delphi is clearly normative. For the most part, these transplants from the original method differ as follows. Rather than speculating about what is probable within a given time frame in the future, the normative Delphi focuses on establishing what is desirable in the form of goals and priorities. The idea of information feedback remains intact. However, the content of that feedback differs. Rather than revealing the dates and probabiliites others assign to future event statements, respondents in the normative Delphi learn the priorities which others assign to goal statements. For example, respondents might be asked to rank the following goal on a scale of highest to lowest priority: "acceptance of teacher trainees without prior educational prerequisites." The information revealed to the panelists in this case would take the form the average rank of the group.

Thus, in principle, the normative Delphi differs from the exploratory Delphi in two ways. First, the substance has to do with what one thinks is desirable, rather than what one thinks is probable. Second, the normative Delphi may be thought of as not strictly temporal. Whereas the exploratory Delphi is always concerned with rather specific future dates, the normative Delphi is not. That is, the panelists usually are not asked to assign a specific date of occurrence to goals, although in some studies rather general time frames are implied such as "over the next decade and one-half." The main function of Delphi, opinion shaping through feedback, is common to both forms.

The normative Delphi clearly serves a different purpose in policy planning. Its use has been to assess the positions constituents and members of an institution (school, school districts, university, etc.) are likely to take on certain goals. It does not necessarily follow that the goals developed in this fashion have any intrinsic worth. Simply because there is agreement on a goal does not assure there is wisdom in its pursuit. (The best example I can think of to illustrate this point is the vote of the United States Congress on the Gulf of Tonkin Resolution.) A committee vote no more assures that an objective is right, than a committee vote insures the future will be what we expect it to be. In this sense consensus is neither a necessary nor sufficient condition for establishing the wisdom of an objective, nor is it a sufficient or necessary condition for establishing the plausibility of a forecast.

Furthermore, deliberately shaping consensus on goals through feedback will have little payoff for policy planning unless certain underlying assumptions are bared in the process. This use of Delphi, like that of forcing consensus about future events, can be argued to be trivial on the same grounds. As presently construed, neither gives much attention to underlying assumptions. In the case of goals, no underlying rationale or motivations are aired to explain why a goal should be accepted as important.[4]

Ranks assigned to goals, based as they likely are on different and sometimes quite naive and even conflicting rationales, are in themselves of little value. It is not enough to simply say a goal is important. One must attempt to give the most powerful justification possible to such an assertion. Otherwise reasonable men have no rationale for rendering a decision, and the intention of such policy instruments, of course, is to aid in the making of decisions.

Family of Forecasting Tools

Used in either of its modes, the Delphi Technique might be characterized as a member of a family of forecasting tools. Let me identify that family. Although Delphi is the focal point of this study, it is one of several "intuitive" forecasting methods. Other methods include futures history analysis, scenario writing, and cross-impact matrices.[5] What distinguishes this family of tools from others? The answer to that question has to do with the explicitness of assumptions in the forecast.

When people mentally construct an image of the future, it is presumed they do so with some model in mind--a particular picture of how things are in the world. The model may be biased or unbiased, valid or invalid; it may be simple or complex. Yet, on the other hand one may not presume the presence of a model at all. We might presume just the opposite. That is, what we observe in forecasting is not the influence of a model but simply a random process, mere guessing, or intuition of the vaguest sort so as to be nothing more than mere speculation.

How can we decide whether a forecast is the product of a random process of guessing, or is the result of a particular view of the way things work? If the models and assumptions which support the forecast are not made explicit, then that question cannot be answered.

Some forecasting methods entail the explication of a model while others do not. In some methods the underlying assumptions, sources of bias and error, degree of reliability, and the validity of inputs are simply unstated. That family of forecasting methods various writers have called "intuitive." Other methods in which the models, assumptions, and biases are stated, we will refer to as "empirical."[6] Delphi and the other methods mentioned above are all examples of intuitive methods. Examples of empirical methods are trend extrapolation and econometric

modeling.[7] The intuitive methods produce results which are entirely subject to unknown bias, but this is not the same as saying their results are incorrect; the point is that there is no way to assess "correctness."

The intuitive tools share some other common properties. They employ collective opinion or subjective judgment as basic inputs to the forecasting process in lieu of quantifiable data. In effect, they operate on the principle that several heads are better than one in making subjective conjectures about the future. It is assumed that experts, within a controlled intuitive process, will make conjectures based upon rational judgment and shared information, rather than merely guessing, and will separate their own hopes and personal motivation from considered judgment in the process. That is, it is assumed that experts are experts because they are objective, take into account new or discrepant information, and construct logically sound deductions about the future based upon a thorough and disciplined understanding of particular phenomena and how they relate. Simply put, the methods are non-data-based and rely on collective expert judgment. (I will return to certain of these assumptions later as empirical matters.)

Furthermore, the forecasts do not begin, as do extrapolations, with a demonstration of how future events grow out of specific present or past conditions. That is, these forecasts are not so much projections as they are quantum leaps into some future time frame in which one is left to find his way backwards to the present.

In summary, intuitive forecasting

(i) employs collective opinion or intuition as basic inputs to the forecasting process;
(ii) does not begin, as do extrapolations, with a demonstration of how future events grow out of specific present or past conditions;

(iii) does not necessarily reveal the models upon which their authors base their opinions nor their sources of inputs to the opinion formulating process;

(iv) thus, reveals little in the way of an understanding about sources of bias, underlying assumptions, and the nature and the validity of inputs.

Explanatory Power

It follows from the above that if one accepts a Delphi forecast as plausible, he does so on the basis of blind faith. The plausibility of a Delphi forecast, as now construed, can be debated only on the basis of the extent of panelist agreement. But agreement alone is not a sufficient condition for arguing that a forecast is plausible and convincing (it is not even a necessary condition).

The nature of the Delphi method ought to be such that certain rather important distinctions could be made about forecasts and their underlying assumptions. For instance, we often fail to distinguish what is desirable from what seems plausible about the future. When we talk about something being desirable in the future, we use such words as "hope" or "goal." When we speak of plausibility, we use such words as "expect," "probability," or "likelihood." There is a fundamental distinction to be made, although often it is not. The purpose for making such a distinction is to separate forecasts of what seems likely--given certain factors--from what we would like to see happen, or like to avoid happening. It is not clear now how one can discriminate between statements in Delphi forecasts that are the products of hope as opposed to those which are products of rational probability estimates. It is clear, however, that hope and desirability interfere with and to a considerable extent influence judgments about future events.

A second fundamental distinction needs to be made. In the absence of actually knowing in detail just what the future will be, one can either guess or judge. The very basis of the Delphi forecasting process is opinion as to when an event is likely to occur. It seems important, in establishing the plausibility of such opinion, that it be supported by rational judgment rather than merely guess-work. Delphi, at present, can render no such distinction because the arguments which support an opinion are not emphasized unless the opinion is contrary to the group norm.

The failure to clarify and share assumptions is a fundamental failure of Delphi forecasts. Studying the future is in effect studying assumptions we hold about the future. Stripping bare the underlying assumptions about the future often reveals that we present no alternatives, our thoughts are based upon very naive and weak arguments, and our judgments are the product of linear thinking. It is therefore crucial that these tools heavily emphasize the explanations upon which the forecast rests. An intuitive forecast which carries with it no explanatory quality may be correct, but it would be trivial. That is the singular weakness of Delphi--in their present form, its forecasts have little substantive explanatory quality.

In order for a forecast to convince men of reason to take some action, on the basis of an argument form presented through a forecast, then the forecast must entail a plausible explanation of what is expected--both why one should be convinced to act, and why, if one failed to act, the consequences foreseen are the most reasonable consequences to expect.

II.

APPLICATIONS OF DELPHI IN EDUCATION

Normative Studies

Delphi has been tried in educational planning on the assumption that almost anyone can forecast the future. As a result, many studies are without the benefit of some considerable understanding of the processes of futures thinking as well as policy thinking, and without the benefit of exact tools that facilitate a more complex consideration of the future than existed twenty years ago. The result so far has been to force attention to the rather naive ability of educators to construct images of the future.

So many studies have emerged that reviews have begun to appear, and now with this report, reviews of reviews. Judd's[8] summary, for example, focuses on the application of Delphi and modified Delphi procedures in university planning. He cites in particular the work of Marvin Adelson, Olaf Helmer, Frederick Bolman, James Jacobson, Arnold Riesman, Samuel Cochran, E. S. Quade, Frederick Cyphert and Walter Gant, and Donald P. Anderson among others. In their work in university planning, Delphi has been applied in one way or another to cost-effectiveness, curriculum and campus planning, university-wide and state-wide educational goals and objectives, and evaluation.

Like the vast majority of offerings on Delphi, Judd's paper tends to be uncritical. He chooses instead to promote the application of Delphi, rather than to dig into its epistemology. Any consideration of the

literature on human information processing and future time perspective, both crucial to an understanding of the Delphi process, is lacking.

One of the earliest uses of Delphi in educational planning was Helmer's Delphi, which was incorporated as part of the 1965 Kettering Project[9] to elicit preference judgments from a panel of experts in education and various fields related to education. The purpose was to compile a list of preferred goals for possible federal funding. Just what value this study had is left in doubt by the experimenters. Helmer concludes, "Although we believe that the compilation of a large number of ideas for possible educational innovations has served a useful purpose, not too much weight should be given to substantive findings resulting from these pilot studies."[10]

Additional Delphi studies are reported as experiments to elicit preference statements from educators, or those with a direct interest in education. Most of these studies are considerably more focused than Helmer's. Cyphert and Gant[11] used Delphi as an opinion questionnaire to elicit preferences from the faculty of the School of Education at the University of Virginia and its clients, regulators, and constituency. Anderson[12] used Delphi in a similar way in Ohio but limited the focus to a county school district. In the Anderson study, statements were obtained from teachers, board members, administrators, and selected educational experts. The statements clustered in two sets: client services and organizational adaptation. Using three Delphi questionnaires, priorities were assigned to the compiled set of goal statements independently using "zero sum" logic.

In both the Virginia study and the Ohio study, most of the change in the priorities occurred after the first modal distribution was reported back to all respondents. Subsequent rounds failed to produce significant changes. The greatest disagreement on particular items in the Virginia study was on preparation of teachers at the graduate level without prior experience, and promoting uniformity of curriculum state-wide. The former

item, preparation of teachers without experience, was ranked among the top ten priorities by the groups as a whole, but lowest by organization leaders and politicians. The latter item, a uniform curriculum, was ranked high by the non-teacher organizations, and low by the university and expert groups.

These education studies differ in principle from the original use of Delphi. In the three studies, respondents were asked to focus on what they would like to see happen, rather than what they considered likely to happen. However, it is unclear as to how that would change the outcome of either type of experiment. It is not possible at the moment precisely to separate Delphi statements which reflect rational judgment from those which are based solely on feelings of desirability. When the task is speculating on the future, just what assumptions underlie one's responses are unclear--unless, of course, those assumptions are specifically and systematically flushed out.

Two other studies which focus on the goals held by various college populations and where Delphi procedures were used to examine values and goals are discussed below.

In a study at Education Testing Service, Norman Uhl[13] investigated goal preferences of off- and on-campus groups. Through questionnaires the two groups were asked to judge the actual importance goals seemed to have at their respective institutions, as opposed to how important those goals should be at their institutions. The on-campus groups, all from Southeastern universities, included students, faculty, academic administrators. The off-campus group included trustees, parents of students, community leaders (politicians, representatives from the business, and religious communities, members of minority groups, and newspaper editors). The polarization on goal preferences which did indeed emerge between these groups after one session in the study disappeared with subsequent feedback.

Among the substantive findings was the none-too-surprising discovery that not one of the groups rated religious orientation high in actual or preferred importance. All groups rated intellectual development high in preference but somewhat lower in implementation. Self-study and planning were also rated high in preference but lower in implementation. The lowest rated preferred goals, other than religion, were national and international service. Perhaps the most surprising finding was the substantially high agreement among groups, particularly when they were often thought to be natural adversaries. Finally, as was the case in both the Cyphert and Gant study and the Anderson study, all of the statistically significant convergence occurred between rounds one and two--after opinion feedback, but before each respondent's defense of his position was fed into the process of judging. But a different effect was seen regarding preferred goals. Significant convergence continued to occur after defense positions were presented.

Dalkey and Rourke[14] investigated the use of Delphi in processing personal judgments about "quality of life" (WOL) as perceived by college students. The factors were generated by the students and refined by a clustering process in which students sorted the factors and then, in effect, pooled their feelings through a process of information feedback.

Among the findings were shifts in cognitive factors, which at first ranked high, but were later moved to lower ranks wnen weighted according to the relevancy they were viewed as holding for quality of life factors. Although the quality of life factors highest in important were affective, i.e., "love, caring, affection," the education factors highest in importance were cognitive, i.e., "ability to learn, learning to learn, reasoning ability, ability to think, critical ability." The education factor seen as most relevant to "love" was the "ability to learn" factor. As an overall educational factor, cognitive skills, when compared to other factors in terms of relevancy, dropped from first to seventh. Self-confidence as an

educational factor moved from eighth to first when relevancy weightings were assigned.

Exploratory Studies

The exploratory Delphi technique has been used in essentially its "pure" form in producing forecasts about the future of education. By "pure" I mean the deliberate use of information feedback to shape the opinion of anonymous judgers about the occurrence of particular future events. As a pilot experiment[15] at the San Diego meeting of the National Conference of Professors of Educational Administration, a Delphi was conducted by staff from the Institute for the Future, Middletown, Connecticut, and the Educational Policy Research Center, Syracuse University Research Corporation. The major purpose was to collect conjectures about prospective developments which might have an impact on educational administration, their probable dates of occurrence, the desirability of such developments, should they occur, and their potential interventions. The study has never been formally reported.

In Canada several studies adorn the growing body of educational Delphi studies. For instance, Berghofer's[16] study was concerned with general education in post-secondary institutions. Clarke and Coutts[17] examined the conjectures of teacher educators.

Berghofer's findings are extensive and beyond the scope of this paper. However, in brief, his study is important for two reasons. First, it is systematic. For example, for the most part, panelists for Delphi studies are selected arbitrarily and somewhat haphazardly. Berghofer's procedure was quite thorough. Second, he modified the feedback procedure in a very significant way. The feedback each panelist received from round to round consisted of the arguments and rationales the other experts

developed in defense of their opinion; dates and probabilities were not fed back to panelists.

Berghofer found that statistically significant differences existed between the final predictions of young and old panelists. Differences of a statistically significant level were also attributed to level of self-appraised competency, level of educational attainment, and organizational position held. In general, the panelists who held highest degrees and also, as it turned out, held educational posts, tended to take the most absolute positions--checking "never" and "perpetual" most frequently of the groups.

Overall, the dates selected by experts for each event after feedback of arguments and rationales tended to shift toward the future. Self-rated appraisal of competency tended to be reduced. Unfortunately, these changes were not statistically treated in the study.

Substantively speaking, the greatest agreement among the panelists clustered in ten problem statements (80% agreement on the year by which the experts thought a majority of people effected by the problems would clearly aware of them). Berghofer reports "A synthesis of this opinion would indicate that the respondents looked forward to a society in which equality of opportunity is emphasized; quality of life is placed above quantity in life; leisure is used creatively; communication skills are stressed; concern is shown for major human problems, and a philosophic basis is sought for social, cultural, economic and medical changes."

Clarke and Coutts found teacher educators generally agreed that teacher candidates would soon have to be skilled in the use of technology, that English usage would be an important criterion for evaluating teacher candidates, and that knowledge and skill in process teaching methods, rather than product methods, would be essential. They also agreed almost unanimously that teaching skills would be required in individualization and group process

as well as in team teaching. The least agreement found in the study centered on rather ambiguous statements about "change," and rather more specific statements about control of teacher education and certification of teachers.

There are other studies reported to date using Delphi in essentially its original form. Hudspeth[18] conducted a Delphi study of perceived vocational education needs in New York State. The population of experts was selected from components of the vocational education system identified by the author as "in," "out," "through" and "external." Particular attention was focused on "electro-mechanical technology and education." Dates and probabilities for each event were subdivided into responses of the four subgroups. All groups received the feedback of the four groups. Respondents were asked to rate each event (presumably should it occur) in terms of value it would have personally and value it would have to society in general. Respondents were also asked to identify each of the four subgroups that had "power" to enhance or inhibit the events. Respondents were finally asked what strategies they would choose to enhance or inhibit the events.

Findings were not treated for statistical significance. The author reports, however, that the majority of events showed convergence but little shift in median date chosen. Events were generally seen as having more value for others than oneself. There was considerable agreement on the subgroups viewed as most influential in altering the occurrence of each event. Strategies for altering events were reported to be "poorly formulated" but tended to fall into five areas: more money for R&D, tax incentives, lobbying, union pressure, increased public awareness.

Doyle and Goodwill[19] conducted a Delphi for Bell of Canada on the future development and utilization of technology. The specific focus was on information systems (computer assisted instruction [CAI], computerized library systems, communication terminals, Audio-Visual retrieval systems). The researchers posed a number of possible developments and requested the panelists to judge their occurrence and also to add to the list.

The substantive findings generally posed a rosy future for educational technology. The experts agreed there would be extensive development and widespread adoption of educational technologies during the late seventies and eighties. Generally it was felt that cultural values would be gradually changing to more openness to innovation, more insistence upon involvement and participation, and more educational practices oriented to the individual.

Delphi was also used to develop long-range forecasts stemming from social indicators in a study conducted by the Institute for the Future, and sponsored primarily by the Educational Policy Research Center at Syracuse University Research Corporation.[20] The areas of concern were: urbanization; international relations; conflict in society and law enforcement; national political structure; values; impact of technology on government and society. The project was part of a larger continuing methodological and substantive study of the future environment in which educational policies enacted in the near term might be expected to have some impact. The study was conceived, not to prepare a detailed description of the future, but instead to examine expectations held by persons well-informed in several domains of the social sciences about the future. The study was intended only to be an initial step and not a final piece of research. The substantive findings from these studies are summarized elsewhere and are beyond the scope of this report.[21]

In brief, a number of difficulties were encountered in the research. First, there was no comprehensive theoretical framework to guide the inquiry. Second, and fundamentally, the social science expectations did not carry the crispness of language and precision of judgment that the more rationalized process of technological change seemed to have in the original uses of Delphi. For instance, just when electric power plants driven by thermonuclear fuel will become widespread is a development controlled by several "knowable" technological factors. The same cannot be said of when alienation and impersonality of urban living will reach

its maximum. Indeed, we do not even know what it means to speak of a "maximum" in this case. Third, the data base available to social science forecasting is shifting and often more unreliable than technological data. For example, data on the percentage of urban minorities is often not valid and its collection a matter of serious controversy. Fourth, even with the best of statistics, judgments in the social domain are subject to considerable variance due to disagreement on the meaning of indicators, and thus forecasts are more likely biased by personal values than may be true of technological forecasts.

Another societal Delphi study, using a format and design very similar to the above, was conducted at the Westrede Institute in Edmonton, Canada.[22] The purpose was to prepare a series of forecasts on social conditions which tend to be important in educational planning. The six topics chosen by the researchers were: changes in value and social goal orientations; the family; leisure and recreation; intercultural relations; politics; and problems and needs of the individual. The purpose, according to the authors, was to be deliberately broad rather than achieve depth. No effort was made to determine possible impacts the forecasts might have on education (although the failure to do that makes the original intent of the study seem rather odd).

The procedures used did not include iterative feedback. Only two questionnaires were used--one requesting a list of forecasts, the other requesting dates, probabilities and rationales.

The substantive results tended toward irony. For instance, the panelists viewed the future as holding much promise for the upgrading of humanist values (personal liberty, social consciousness, self-respect, etc.), but at the same time predicted seriously widening divisions between young and old, English and French, red and white, rich and poor, East and West. The panelists expected the education system will be more responsive

to the needs of students. Yet, they also expected disaffection to increase, and felt that nothing short of radical overhauling of the fundamental structures and processes of education would be necessary. More specifically, the panelists felt conflict in higher education would worsen between student and institution. But they also expected authoritarianism to decrease, student participation in decision-making to increase, and curriculum reform to lean toward creativity, personal relationships, change process, leisure time. They also expected great increases in demand for continuing education.

Although not clearly reported, the greatest amount of disagreement in the report seemed to be in human relations areas: law and order, violence, and alienation. The researchers summarize the salient findings in the following themes: aspirations and demands for social reform will outstrip actual reforms; society is in transition; many institutions are experiencing obsolescence; individuality and personal freedom will be upgraded; and individuals will be frustrated. ("The strongest theme among these forecasts pertains to the frustration of the individual.")

In short, the forecasts anticipate the best of times and the worst of times.

Finally, Delphi has been modified and linked together with other tools, not for the purpose of producing intuitive forecasts, but for the purpose of modifying the awareness, assumptions, and skills of the persons making the forecasts. For example, Sandow[23] constructed a simulation exercise which links together in a logical flow of activities the basic principles of Delphi, Cross-Impact Matrix, scenario writing, and analysis of future histories.

There have been a number of other "first step" efforts elsewhere to recast forecasting tools such as Delphi into teaching tools. These efforts are largely unreported to date. The "Ghetto 1984" game developed by

Professor Jose Villegas at Cornell University[24] bears noting, as does the Delphi Exploration Game developed at the University of Illinois.[25]

In the University of Illinois project, initiated by Professor Charles E. Osgood, Delphi was used to create a computerized gaming device called Delphi Exploration. The general pattern of the game followed Future, a parlor-type game developed by Olaf Helmer and Theodore Gordon. Statements from prior Delphi research were used in the computer game. In addition, the cross-impact matrix has been added.[26] In Delphi Exploration the players make investments in one set of future events in an attempt to move undesirable developments toward 0 percent probability while moving desirable developments toward 100 percent probability. In the Delphi II program now under development, the player will be able to work through time from the present to some point in the future. In Delphi I, the operating program, the player simply tries to build what he considers to be a desirable world in the year 2000. It is the process through which players must go in Delphi Exploration that seems to be its objective as a teaching device.

Some Criticisms

There are several weaknesses inherent in the Delphi methodology as construed in these studies. First, there is the failure to distinguish between assertions, which may or may not be right, and their more important underlying explanations and assumptions which could be judged as reasonable or unreasonable. Second, it is assumed that consensus and plausibility are somehow connected. That is, if people agree on something, it must be right. We have argued elsewhere that in principle consensus is neither a necessary nor sufficient condition for saying something is plausible. Furthermore, consensus clearly does not mean that rational judgment was exercised in the process. It has been empirically demonstrated that agreement can be achieved even when agreement clearly runs counter to logic or observed reality.[27] Third, the present applications of Delphi seem to

represent "establishment futurology." The first of these weaknesses was discussed earlier and the second will be discussed in more detail in the next section. Let us take a moment here to discuss the third area of criticism--the tendency of Delphi to become an instrument of establishment futurology.

This was given some attention in "An Interim Report on the Alberta Delphi Interaction Studies." Unfortunately, the criticisms did not survive in the final report.[28] However, we want to give it an airing here, together with some embellishments of our own.

The Delphi studies reported above tend to be descriptive rather than explanatory. They generally are surprise-free, suggesting no major discontinuities and implying that current trends will continue, perhaps more sharply, perhaps not. They are, in appearance only, value-neutral; however, under the surface, they clearly present the views of an incipient bureaucracy. For instance, there is a failure to recognize the difference between "schooling" and learning"; this leads to the erroneous conclusions that learning occurs mostly or even exclusively in schools, and that when the demand for learning increases, schooling must also expand.

There is a serious confusion in the way problems are defined. The confusion is carried over into the future. For example, there is a persistent failure to distinguish between what schools do to individuals and what schools do about individual differences. Consequently, numerous forecasts confuse the problems of self-expression, alienation, and individuality among youth with institutional proposals such as IPI.

Finally, not one of the studies reported here includes the views of the radical political left. Establishment futurology is entirely characterized by the talk of those who really are satisfied with their particular positions and roles and status, although in that talk, certain popular

metaphors and euphemisms of change are generously allowed. Those who are really dissatisfied, those whose ideas do not fit, will reject this mode of futures research, and probably would not participate in a Delphi even if asked--which is unlikely.

The vast number of Delphis which have been run in various educational institutions suggest that there is something to the argument that Delphi has been seized as an instrument of establishment futurology. The educational Delphis are in no way startling or sensational. That is obvious to the most casual observer. There is a serious sterility in the process of summarizing mass information into numerous narrowly terse statements. There is a serious absence of any effort to probe beneath the surface for explanations. In their make-up, Delphi panels cater to the power structure, not the disenfranchised. Furthermore, the Delphi studies reviewed here suffer from technical limitations imposed by the methodology. Topics selected for consideration depend on the subjective judgment of the experimenter or his panel. Specific content is particularly subject to experimenter bias because of the necessity to collate and summarize responses. Choice of alternative response forms us subjective, and generally no provision is made for estimating the effects of greater or lesser alternatives. There is no provision in the studies to check on the effect of wording, order of items or other devices that may influence the predictability of events.

Delphi studies ought to be received critically, evaluated thoroughly and taken seriously--if they are actually believed to be an input to planning. Now, unfortunately, that is not the case.

III.

EXAMINING THE RESEARCH BACKGROUND ON DELPHI

Future Cognition: Two Traditions of Research

There are two distinct bodies of research literature relevant to Delphi. The one literature is on the deliberate shaping of opinion through information feedback (which is related tangentially to the vast literature on response set and personality). The other is a literature about what I have begun to call "future cognition," the study of human thought and the future.

A detailed review of the literature on these topics is beyond the scope of this report. However, below is a brief excursion through one salient aspect of the future cognition literature. Other studies from other aspects of relevant research will be discussed where appropriate.

There have been two notable trends during the last two decades in the study of human thought and the future. One trend has been the investigation of future time perspective and personality traits. This research has been reported primarily in the literature on abnormal psychology. Investigations have been concerned with social relationships, time perspective, and personality. Studies range from theoretical and experimental reports to speculative papers.

The second trend appears in the literature of forecasting, public opinion, and information science. These studies have been concerned primarily with accuracy of forecasters. Studies range from experimental to

applied research, but differ fundamentally from the psychological studies. The central thrust has been methodological, i.e., oriented toward research and development. The research goals have been not so much to explain human behavior on the basis of some theoretical construct, but to improve the forecasting techniques.

The relation of these two trends could best be described as coincidental. They begin with a common problem, namely, the relation between human thought and the future, but they have developed in isolation from one another. There is little evidence of cross fertilization of data, ideas, constructs, or theories. An exhaustive check of the documentation shows two disparate and neatly isolated traditions of research--the experimenters in the one tradition not citing studies in the other. Yet, within each tradition there is considerable continuity.

While the specific research itself on Delphi is meager, the evaluation and reporting of even what little there is remains an arena of neglect. Few, if any, of the serious authors of Delphi reports (and we include all of those noted in this report) have investigated the two traditions of research in any fashion resembling an evaluative approach. This neglect is the primary reason for detailing the studies that follow.

There is a clear difference in the focus of the two traditions of studies. On the one hand, experimenters in the time perspective studies have attended to personality traits without systematically testing the relation of such traits to the particulars of the experimental condition, i.e., level of ambiguity, level of uncertainty, degree of complexity, level of abstraction, the nature of feedback and types of reinforcement--all of which, it can be argued, are present to one degree or another in each of the experiments.[29]

On the other hand, experimenters in the forecasting studies generally tended to focus on procedures, but not on the interrelationship between

procedures and personality traits. The notable exceptions are the McGregor study,[30] and the more recent studies of Campbell,[31] Weaver,[32] and Waldron.[33] For the most part forecasting studies have focused on the variance which could be accounted for by feedback of certain forms of information, nature of reinforcement, task complexity and the like.

The Early Studies: Forecasting and Forecasters

Nathan Israeli was one of the pioneers in subjecting forecasting to empirical test.[34] His work, however, has been criticized for containing several methodological and conceptual defects.[35] Israeli's work is nonetheless mentioned here because several of the procedures he used or proposed for studying future cognition are now being used in Delphi forecasting techniques.

In one study,[36] for example, subjects made a series of qualitative and quantitative judgments about future events. Subjects were asked to respond to future events in the following ways: (a) set a date for their occurrence, (b) select the most probable development among alternatives, for a specific future date, (c) select the most probable outcome from among alternatives for a given situation at a stated future time. These three tasks have been further experimented with by McGregor,[37] Kaplan,[38] and Helmer.[39] In other studies Israeli experimented with wishes of college students regarding "improbable" future events in an effort to explore conflict between wishes and reality[40] and student emotionality toward past, present, and future.[41]

Israeli developed a series of ten experimental designs in all to be conducted over a period of time.[42] The results of these experiments unfortunately were incompletely reported in the literature. However, the research designs were important in terms of questions, assumptions and rationales. Among the ten designs are the following which have been given

attention by experimenters in one way or another over the last three decades. The designs also raise some very important questions not explored to date.

One of the designs would have tested the degree to which known information affects a prediction. This was later tested by McGregor. Israeli was concerned about the degree to which projections would follow logarithmic curves when subjects are or are not given past information about the event. A second design would have tested the degree to which "dogmatic and more liberal or imaginative" individuals would perceive remote and near futures. Roberts[43] and Bonier[44] later explored this question, but with results Bonier interpreted as contradictory to theory. This question is important because of Israeli's assumption that extension into the future is accompanied by increasing variability in thinking about different situations, and his assumption that dogmatism is perhaps relevant to such increased variability.

A third design would have subjects rank the importance of certain eminent authorities at various times in the future as they might be ranked by those living at the time. Still another design would have asked subjects to name a period of time in which catastrophes occurred by certain areas, e.g., chemical, sociological, etc., and to name a factor contributing to the catastrophe. A fifth experimental design dealt with probability or certainty felt by the subject that an event would occur in a given period of the future. As a particular aspect of probability, Israeli suggested using the catastrophes elicited in earlier designs for a set of stimuli to which subjects would assign probabilities. The assignment of probabilities to the occurrence of future events is a major source of data for the Delphi technique and the related cross-impact matrices.

In short, the Israeli studies are significant not so much for elegance of findings, but for the very basic questions he raised about human thought and the future. These questions have continued to be explored for more than thirty years.

Two important studies followed the Israeli research. McGregor[45] dealt with the problems of "predeterminers" of prediction, i.e., attitudes, wishes and beliefs, together with "objective conditions that have been present in the immediate past in the environment of the predictor." His assumption was that these factors taken as a whole determine the premise underlying predictions.

McGregor found that the highest probabilities of occurrence for an event, regardless of the respondent's feelings of desirability, were registered in situations thought by the respondent to be unambiguous. McGregor concluded that predicting occurrence of events seemed to be coerced by a reduction in the ambiguity of the situation surrounding the event. McGregor also found that the greater the ambiguity of the situation and the importance of the event to the predictor, the closer the prediction corresponded to attitude. Familiarity was not related to differences in 8 out of 9 predictions. Although the experts were "much better informed on the average," their predictions did not differ significantly from those of students when their attitudes were roughly the same.

The McGregor experiment is particularly significant because he began to explore the interrelationship between dispositional factors (attitude, etc.) and the nature of the judgmental conditions (ambiguity, uncertainty, subjective probability).

A study similar to McGregor's work was conducted by Cantril.[46] Cantril was interested in exploring to some degree the following kinds of questions. Can predictions of the timing of an event be made with as much certainty as predictions regarding the actual outcome of an event? Are there differences in the accuracy with which local and geographically distant events are predicted? Are there differences in the certainty with which immediately likely and distant (near and remote) events are predicted? What is the comparative accuracy of predictions of individuals versus groups? Are "men of affairs" more certain of their judgments than academicians? What is the

effect of the attitudes of predictions? What are the circumstances under which attitudes contradict predictions?

Cantril found that all of the respondents were more certain that an event would occur some time than they were about the exact date of occurrence. He found that attitudes, as determined by an eight item survey, tended to be related to affirmative predictions of events in the direction of the attitude, e.g., socialists tended to affirm such "socialist events" as federal control of electric power.

In addition, Cantril found that events which appeared to lack relevant facts for a predictive base were judged with little certainty, and that academics were less certain of outcomes of events than "men of affairs," e.g., bankers, insurance executives, newspaper editors, etc.

The McGregor and Cantril studies differed in the following respects:

(i) McGregor attempted to relate personality attributes to accuracy of event probabilities. Cantril could not control for accuracy because only 2 of 15 events in his study actually occurred during the study.

(ii) Cantril explored the differences in probabilities assigned to date of occurrence as compared to actual likelihood of ever occurring. McGregor did not explore this question.

(iii) Cantril explored the question of whether academics and men of affairs differed on certainty of predictions and whether certainty was related to probability. McGregor examined the differences between academics and students.

It should be noted that neither the McGregor study nor the Cantril study explored the effects of information feedback.

More Recent Studies on Forecasters and Forecasting

An experiment similar to the two above, was conducted by Abraham Kaplan.[47] Kaplan raised questions about three basic areas of prediction: evaluation, improvement, and appraisal. With regard to evaluation, Kaplan was concerned with these questions. How successfully can predictions of social and technological events be made? Do social and technological events differ with regard to stability of predictions? How precisely can predictions be made? Are there differences in predictions of near and distant future events?

Kaplan defined three basic research problems in the area of improvement: improvement of prediction reliability by taking the mean estimates of the probabilities of events; improvement of accuracy by weighting the probabilities according to prior performance of experts; improvement by collective group predictions as compared to a number of predictors working independently.

Regarding appraisal, the essential problem was viewed as one of "specifying subpopulations of predictions in which the probability of success remains relatively stable." Specifically the stability question was related to confidence and precision of estimate. The question was whether predictions made with high confidence are more likely to be successful than those made with less confidence.

A twenty-week limit was set on each event to be judged in the study. The event could be confirmed or not within twenty weeks. In each questionnaire the respondent was given four exclusive predictions to which he was asked to assign probabilities of occurrence from 0 to 100. Values for the four alternatives were to sum 100. In addition, space was provided for an open-ended statement of "Basis for Your Judgment."

The questionnaires were distributed weekly for 13 weeks to all respondents. One-half of the predictors worked together on each new set of predictions in quartets split as follows: (a) an independent group answered individually as usual; (b) a cooperative group discussed the questions, but answered them individually; (c) a joint group discussed the questions, came to a collective decision and gave one answer for the entire group. The participants rotated among the three groups.

Kaplan found that the relatively near future was more accurately predicted than the relatively distant future. Prediction success varied inversely with its scope in time. Five months was the longest interval of time considered.

The entire group on all questions had a success mean of 53 percent. That is, in 53 percent of the cases where highest values were assigned, those cases were verified. Random success would have been 25 percent. However, predictors who were often right were scarcely more definite than consistently wrong predictors, definiteness being the degree of success above 25 percent.

Natural science events were no more successfully predicted than social events--despite the fact that most predictors were not social scientists. However, Kaplan found that social events were predicted with more confidence than science events.

Knowledge of specific events was not related to successful predictions of specific events in the study. Kaplan also found that a statistical averaging of independently made forecasts yielded a success rate equal to group formulated predictions. However, joint group efforts and cooperative group efforts (discussion and the independent predictions) were superior

to predictions by the same individuals when not participating in groups.

Finally, Kaplan found that one's justification of the basis of his judgment was related to successful prediction--justification was defined as a statement of logical warrant for prediction. Examples were factual elaborations of details of question and answer, evidence of specific empirical generalizations, hypotheses about motivations of predicted behavior, analysis of time required for the event to occur. "Guesses" (suspected by Kaplan to be systematic or "educated") as the stated basis of prediction were successful in 40 percent of the cases--significantly better than chance.

Kaplan's study is quite significant in the tradition of research on prediction behavior. However, a discussion would not be complete without citing some of the disclaimers in his research.

First, topics selected for the questionnaires depended on the subjective judgment of the experimenters with regard to such factors as likelihood of occurrence of an event within the five month interval, and the intrinsic difficulty (level of complexity) of each prediction. The specific content of the questions was subject to experimenter bias. Second, choice of alternatives was subjective and the study provided no basis for estimating the effects of greater or lesser numbers of alternatives, or allowing the predictors to specify alternatives in an open-ended fashion. Third, there was no opportunity to check on effect of wording, order, or other devices on predictability of events. Fourth, the short time span in the study had the effect of tending to force selection of questions for the questionnaire from among potentially rapidly changing events; the unexpected consequence of this was to enhance predictability of certain items on the basis of obvious forecasts of "No Change." The first three of these limitations would apply in general to all Delphi studies.

Feedback Studies

Norman Dalkey has for several years engaged in a series of studies of group formulated opinion at the Rand Corporation.[48] His interest parallels that of Kaplan, and also that of the earlier McGregor and Cantril studies. Dalkey is particularly interested in the question of improved (more accurate) group judgments through the use of controlled feedback.

In Dalkey's experiments, the questions have typically been drawn from almanacs and therefore the answers are of a factual nature which can be confirmed or disconfirmed. They are in that sense atemporal. The questions do not in themselves demand any consideration of the future. Whether estimates made under such circumstances bear any relevance to Delphi forecasting is a matter still untested. The judgmental tasks inherent in forecasting might be presumed to differ on logical grounds and may differ on psychological grounds as well.

Two basic problems were investigated in the Dalkey studies: comparison of face-to-face discussion with controlled feedback and improvement of group estimates using an iterative form of information feedback.

In general, Dalkey found, "more often than not," that face-to-face discussion tended to make estimates of the group less accurate, whereas controlled anonymous feedback made the group estimate more accurate. Specifically, he found that the median response of the questionnaire group was more accurate in 13 cases out of 20 and the discussion group more accurate in 7 cases. The result was not statistically significant. In an a posteriori experiment using smaller groups but giving them anonymous feedback from a Delphi questionnaire prior to their forming groups, Dalkey found that discussion after the first round produced more accurate estimates, but further discussion also produced more inaccurate answers.

Dalkey tested the widespread belief that, in the controlled feedback process, group "agreement" or convergency means answers are more likely to be correct than if the group's response remains widespread. The correlation between standard deviation (spread) and accuracy produced, in Dalkey's words, a "disappointing result" (statistically significant, "but not high enough to be interesting"). Dalkey, in comparing estimates to a randomized set of answers, found that differences were "heavily masked by chance" on the first round.

Dalkey also found that repeating the feedback from round to round had the effect of closing the spread and also improving the medians of some answers while reducing the accuracy of others. For about 64 percent of the changed estimates, the median improved in accuracy; but for 36 percent, the median became less accurate. He found as well that the respondents closest to the median on the first round were the most accurate and also less likely to change. After iteration the swing group became more accurate as the median shifted in the majority of cases toward the "true" answer.

Finally, it is clear that the group norm is much stronger than the effect of the "true" answer. That is, the convergence is consistently toward the group norm (median) independently of whether the norm moves toward the "true" answer. This, of course, leaves unanswered two fundamental questions. Under what conditions will people readjust their estimates toward the norm, and what are the effects of personality in such a process? Is there even the slightest relation between convergency and "correctness" in estimates and, if so, how could one explain it?

In conclusion Dalkey[49] cited three basic findings on prediction that have emerged from the experiments at RAND: pronounced convergence of opinion occurs after feedback; the major part of convergence takes place between the first and second rounds; and, in cases where accuracy could be checked, the accuracy of group responses increases with feedback. Finally, Dalkey reported that considerable variance existed in performance on

different questions in the experiments. Split-half reliability on questionnaires ranged from .4 to .6. In Dalkey's words, not high enough to "measure with."

Brown[50] also examined the question of accuracy in prediction. The study did not attempt to explain why accuracy should or should not occur, or why accuracy might vary in a given population. Like the Dalkey experiments Brown's study did not deal with future events, but instead also used almanac-type questions. In Brown's study twenty-three RAND researchers were used as subjects. Twenty questions were submitted to them. Eighteen of the questions varied in content but could be answered with factual information from the World Almanac. The remaining two were mathematical questions that could be computed but with some difficulty. Each respondent self-rated his confidence on each estimate. Questions and responses were submitted to respondents over four rounds. Each round requested revision of an answer and, if the answer were outside the interquartile range, to state reasons for divergency.

Brown found that medians tended to move more closely to the correct answer over succeeding rounds, but the interquartile ranges converged away from the correct answer, viz., as the range of estimates decreased, the correct answer was no longer included in any of the middle 50% of the estimates. The "ball park" answers (within 25 percent of correctness) increased from 21 percent to 38 percent over four rounds as calculated from the medians. Quartiles containing the true answer decreased from 13 out of a possible 20 to 7 out of 20 over four rounds. Brown also found that the sub-group estimaters who rated themselves highest in confidence had collectively better median success than the average.

Personality Influence

Three recent studies have investigated the effects of personality influence on the outcomes of Delphi forecasting. Campbell[51] compared the effectiveness of the Delphi questionnaire technique against group-discussed forecasts. The forecasts were concerned with specific economic indicators such as GNP. The projections were made three months in advance of confirming data.

Campbell found that the Delphi group estimates decreased more in interquartile range than the discussion group estimates. However, the convergence that occurred in the Delphi process tended to exclude the "correct" answer, and the exclusion process increased over rounds. Campbell also found that individual estimates in the Delphi experimental group were more accurate, in the sense they deviated less from the correct answer, than individual estimates in the discussion groups. His data also reveal that the Delphi sample, as a group, was not more accurate to begin with, but tended to improve over four rounds. Campbell found that, in general, as a group, Delphi forecasts were more accurate than the discussion group forecasts.

Campbell also found that self-confidence (self-rated) tended to be related to accuracy, but he states that "selecting the most self-confident members of a group . . . was not an effective means of identifying the most accurate forecaster" (p. 112).

Campbell found that in the discussion groups frequency of participation, as perceived by the group, tended to be related to the groups' perception of influence and competence and even though the substance of subsequent forecasting tasks changed, the same people were perceived to be most influential and competent. The data were inconclusive in determining whether "influentials" were also more accurate as a subgroup of forecasters.

In about half of the cases they were more successful, but in an equal number of instances they were not. The data did suggest that accuracy of the group tended to be a function of the accuracy of the most influential forecasters.

Finally, and perhaps most significantly, two measures of personality traits were predictive of certain behaviors--both in the discussion and in the Delphi experimental groups. Campbell found that participants with "inclusion" and "affection" needs (FIRO-B scale) tended to be persuaded to change more frequently in the discussion groups. He also found that Delphi was not immune to such conformity-induced behavior. In the Delphi experimental groups; participants with high inclusion and affection needs accounted for a part of the convergence. They were significantly more conformist as a subgroup than others.

EPRC Research

Research conducted at the Educational Policy Research Center at Syracuse has been primarily an investigation of how human information processing (conceptual level) is related to prediction. It was assumed in our study that the spread in estimates made by forecasters could be predicted by their conceptual level and that their estimates would chance in predictable ways under different conditions in the experiment. Our study was also concerned with objectivity as an influence on Delphi outcomes, and whether Delphi is immune to different propensities of forecasters to conform.

With graduate students in the field of education, Weaver[52] found that conceptual level is clearly related to the outcomes of Delphi forecasts. Deciding how far apart to place earliest and latest dates of expected occurrence of future events was a task distinguished by conceptual level.

When information cues were given to help forecasters decide when an event might occur, concrete persons (low conceptual level) narrowed the

distance between their earliest and latest dates. However, when estimating the occurrence of future dates was open-ended, concrete respondents widened their estimates greatly. This effect is particularly significant when it is considered that abstract persons (high conceptual level) did not differ significantly from treatment to treatment.

It was also found that the dates assigned by abstract persons correlated significantly with their self-rated feeling of "desirability" regarding the events. However, the correlation was not significant for concrete persons. The complex thinker seems to be no more "objective" (and perhaps less so) than concrete persons in assigning dates to future events. This, of course, is contrary to expectations.

Waldron's dissertation,[53] which primarily reports a replication of the basic work in the above study, dealt with the propensity of concrete persons to converge their estimates over rounds of the Delphi. In Weaver's study described above there was no iteration of rounds. Subjects in Weaver's study were randomly assigned to treatments which approximated three rounds of Delphi, but no repeated measures were employed. In Waldron's study, high and low conceptually complex persons were subjected to three rounds of Delphi over time. Waldron reconfirmed that feedback of dates narrowed the estimation ranges of concrete subjects but did not significantly affect the highs. He found also that lows had a greater propensity to change estimates across rounds, and to change their estimates toward the norm of the controlled feedback.

Notes on the Recent Research

Although there appear to be some similarities, these studies differ somewhat. Campbell was interested in determining whether Delphi is superior to "uncontrolled discussion" in producing more accurate forecasts. Weaver's and Waldron's studies were not concerned with accuracy. Therefore, while

the three studies confronted subjects with future events, Campbell's events were of an entirely different nature. They could be quantified and covered only a short time span. Moreover, in Campbell's design subjects were not asked when an event would occur, but instead were asked to make projections, e.g., what the GNP would be three months hence. In both Waldron's study and in Weaver's study, subjects were asked to assign an earliest possible and latest possible future date to rather general events, e.g., "widespread use of sophisticated teaching machines." The events could be viewed as extending several years into the future.

Weaver's research was concerned with the differential effects of such experimental conditions as fixed alternatives on persons having different conceptual levels. Campbell's experimental treatments were designed to find out whether the Delphi method would produce more accurate results on short-term projections than would a committee.

All three studies did attempt to explain why convergence occurs by demonstrating that people with certain personality traits conform to an information norm while others--with different traits--do not. Campbell's study is significant in demonstrating that in <u>both</u> group discussion and in the Delphi process personal needs for inclusion and affection account for a part of the convergence. Both Weaver's data and Waldron's data suggest that conceptual level accounts for considerable convergence. It is probably reasonable to assume that the traits measured in all three studies are related, and are predictive of conforming behavior under conditions where specific knowledge of the "correct" answer is lacking, regardless of whether the content of the experimental question is the past, present, or future. That assumption, of course, needs to be tested.

It should be noted that the experiments by Dalkey[54] and Brown[55] should be interpreted carefully in regard to the prior research of Cantril,[56] Kaplan,[57] and McGregor,[58] and the more recent studies of Weaver,[59] Waldron,[60] and Campbell.[61] The findings of Dalkey and Brown are not necessarily

relevant to any temporal consideration whatever. Just exactly what judgmental tasks are involved in these studies is not clear; they may include things such as "ranking," "estimating," "computing," and so forth. It may well be that such judgments are in fact closely related to judgments made about the future. This, however, is an unresearched assumption.

It is also important to note that, except for Waldron's and Weaver's, all of the above research is concerned with "accuracy" of judgment. Although this is a question certainly worthy of consideration, not enough of these studies raise significant questions about the effects of personality biases (attitudes, values, beliefs, etc.) on the outcome of Delphi forecasts. The need for further research into the effects of personality has certainly been demonstrated and some directions are beginning to be clarified.

What we know about how the mind constructs images of the future remains rather puny, but the fundamental assumptions which are generally held about Delphi seem questionable. For instance, the Delphi technique was created to prevent professional status and high position from forcing judgments in certain directions when panels of experts met. The intention was to assure that through questionnaires, changes in estimates would reflect rational judgment, and therefore not be subject to social psychological factors. Empirical evidence tends to show the naivete of such an assumption. Experimental evidence[62] clearly demonstrates that using a questionnaire technique to generate information feedback does not eliminate the effects on conformity that one observes under group pressure. Those persons who tend to conform under group pressure seem to do so even when the norm which attracts them is the statistical averaging opinion from a questionnaire. Furthermore, the conformist (in both types of conditions) tends to be more submissive, more anxious, more authoritarian, less intelligent, less theoretical, less realistic, and more emotionally reactive. The conclusions from this literature generally tend to emphasize the role of motivational systems in explaining conformity. The differences

in the way people seek and use information feedback clearly follows personality patterns--regardless of whether conformity is induced by group pressure or questionnaires.

Three independently conducted studies suggest that within the Delphi procedure individuals who "swing" in from wide ranges to more narrow ranges do so less on the basis of rational argument, examination of evidence, or review of assumptions, than because decision-making strategies of certain persons are subject to change as the task is perceived to be less ambiguous, and on account of certain personality factors such as fundamental needs and integrative complexity. These findings, of course, are not unexpected, and generally support the studies of several other investigators. The propensity to conform might be distinguished as follows in the literature. Conformity to a group norm of unanimous peers who have expressed a judgment which is in obvious contradiction to logic and reason ought to be and is associated with personal and motivational attributes: timidity, deference to others, central needs, needs for approval. On the other hand, conformity to unanimous norms in ambiguous situations which defy logical and reasonable argument should be not only associated with the above attributes, but also with informational-handling conventions such as persistence in seeking closure and external locus of control. Indeed in the EPRC experiments we found just that.

It also seems clear that subjective judgments of even very complex or abstract thinkers may be considerably influenced by their feelings of desirability regarding the future events in question. The assumption that experts, who may be presumed to be complex thinkers, bring to bear "cool analysis" in their judgments about the future, is questionable in light of our findings.[63]

Still focusing for the moment on process, just what do we know about how people think about the future? From the research reviewed, in earlier papers[64] I have drawn the following summary observations.

The psychological studies of future perspective and personality traits strongly suggest that concepts and perceptions held about self and others are interrelated and reflective of thoughts about the future. Conceptual level, alienation, anxiety, social deviancy, emotional instability, and schizophrenia--all powerful indicators of particular ways of perceiving and relating to society--impinge upon one's future cognition. Numerous studies showed that these indicators were sufficiently strong to distinguish perceptions about the future, particularly when such perceptions involve estimating how long something would take or involve foreseeing some state or states of affairs.

It follows that persons with different kinds of "self structures"[65] (needs, attitudes, beliefs, etc.) would hold different perceptions about the present as well as the future, and thus produce different kinds of forecasts about the future. This statement appears to be rather evident. How to shape it into a researchable set of questions is not as evident because exogenous variables also impinge upon judgments. For instance, the phrasing or complexity of a question, or the influence of a group norm, even though it may be anonymous, influences the judgments of certain people. Whether or not the judgmental task is vague or uncertain, or is perceived to be vague or uncertain, may also influence particular people to a considerable degree.

Research questions on forecasting methods must begin to reflect some consideration of the <u>interaction</u> between dispositional factors and the conditions in the experiment. Among the more important questions are how do differences in judgments about the future reflect differences in the self-structure of the people who make the judgments? And consequently, how will differences in estimates be shaped by exogenous variables such as complexity and ambiguity of the task? The failure to consider these questions is a persistent weakness of most Delphi studies to date.

Some Conclusions

Delphi, like the future it was intended to foretell, has not turned out to be what we expected. It displays certain fundamental weaknesses in its present form as a forecasting tool. Briefly, they have to do with interpreting the significance of convergence of opinion under the conditions imposed by Delphi. The observation that people tend to shift their estimates toward a group norm under conditions of iteration is, on the basis of several controlled experiments with Delphi, a consistent and sound observation. There is some very meager evidence which suggests that compression of estimates over rounds produces a final consensus closer to the "true" answer (when the consensus is taken as a median of the spread of estimates). This finding, however, is based upon evidence collected from very short-term predictions in the economic domain, and from experiments with almanac-type questions. Just how accurately the findings can be generalized to Delphis which cover a 30-year extension into the future is unknown. Moreover, to make such a generalization is irrelevant to an understanding of plausibility as discussed earlier. Yet interpreting the social psychological significance of the convergency that does occur with such opinion is important in understanding how the mind processes information about the future. Once we can understand more clearly how the mind formulates images of the future, we will be in a better position to improve upon the process of constructing rational and plausible forecasts.

Any consideration of the future of education should attempt to clarify what we can reasonably expect to make happen or not expect to make happen. Rather than a focus on "accuracy," the focus might better be on "plausibility" or reasonableness of forecasts. In that sense Delphi at present comes up short because there is little emphasis on the grounds or arguments which might convince policy makers of the forecasts' reasonableness. There are insufficient procedures to distinguish hope from likelihood. Delphi at present can render no rigorous distinction between reasonable judgment and

mere guessing; nor does it distinguish clearly priority and value statements from rational arguments, nor feelings of confidence and desirability from statements of probability.

Of equally great importance, however, our research also leads us to conclude that Delphi, in combination with other tools, is a very potent device for teaching people to think about the future of education in much more complex ways than they ordinarily would. When we understand this use of Delphi, we may find that it is a useful instrument for something more important than what it was designed for, viz., a general teaching strategy. What this means is that initially the way to get educators to make better decisions--decisions which account for alternative future consequences--is to enhance their capacity to think in complex ways about the future, and Delphi seems ideally suited to such a purpose. Indeed, educators may find in Delphi and other forecasting tools a better pedagogy.

IV.

MODIFYING DELPHI

Rationale

There are two particular points of focus in modifying Delphi. They are related. For one thing, in the reporting of a Delphi forecast the propositions of its authors are intended to be persuasive, although they carry no particular means for making them persuasive. Secondly, convergence of estimates from round to round is often confused with reasoned agreement. Let us develop this last point a little further.

When feedback consists merely of a distribution of dates or ranks, convergence which occurs carries no rational justification. This is not to say there are no reasoned changes connected with convergence; we are merely pointing out whatever the reasons, they are generally unknown. Nor is this to say that respondents, in finding others disagree, might not rethink their own position and alter it to more closely approximate the norm. But to make that claim is not the same as saying the outcome of the Delphi forecast was influenced by the singular or combined weighing of the arguments put forth by members of the Delphi panel. Those arguments simply remain hidden.

Thus, our purpose in modifying Delphi is to force attention away from its propositions, whatever they are, and to focus on the elaboration of underlying assumptions and explanations. Whatever the outcome of a Delphi exercise, we want to be able to conclude it was influenced by the combined weight of the arguments presented. In this sense, we are talking more of a pedagogy than forecasting; therefore, we are not so much interested in the claims people make about the future as how they support those claims and what they learn from each other in the process.

In this view, what are the basic aspects of Delphi that need changing? First, there must be a redirecting of activity in Delphi studies away from mere description of specific events to explaining why such descriptions are reasonable, and why they, rather than others, are the most significant considerations to think about.

Second, there must be a process whereby authors of forecasts not only explain why their forecasts are to be taken as reasonable but also why they might be expected to occur sooner rather than later. Third, by using those kinds of explanations as feedback, rather than using a statistical averaging of dates or ranks, one would reveal for the total group what each member states he expects and how that expectation is justified.

These changes are crucial if Delphi is to survive the first blush of enthusiasm its proponents have generated. However, such changes in Delphi do not assure forecasts will be more accurate, nor that goals will be right. They simply provide that the reasons men give for such assurances are stated.

While the concept of Delphi forecasting needs to be changed, it also needs to be expanded. It is not clear now how the outcome of a Delphi study is influenced by the panelists' considerations of how events are interrelated. Nor is it clear how the forecast is shaped by consideration

of intervening events and long-term trends not specifically mentioned by the panel. Mechanisms for bringing out these considerations must be added to Delphi. However, simply to add the cross-impact matrix (CIM)[66] in its present form is insufficient. Indeed, to do that compounds the problem. Why?

The Delphi method alone raises enough questions of a sufficiently high order of inference to demand some considerable explanation. If one goes beyond that level of inference and asks not only what is the probability of X occurring at a given date, but asks whether as a result Y will occur, then the complexity as well as the importance of the explanation becomes paramount. Delphi and Cross-Impact Matrix techniques raise questions demanding complex levels of inference and explanation. Yet, they lack the necessary mechanisms to elicit and present explanations, arguments, or underlying assumptions that would allow reasoned evaluation of their results. Therefore, their results cannot be assessed as valid or invalid, plausible or implausible. Because of that, Delphi and CIM lack potency as policy instruments. They simply are not convincing in their present form.

To my knowledge, these needed changes were recommended for the first time by this writer in spring, 1969:

"It may be more consistent with information processing theory to eliminate consensus forcing procedures altogether from forecasting and substitute feedback consisting of (a) assumptions, (b) causal factors, (c) evidence, or (d) theoretical bases. After several rounds of exchanging bases of judgments, rather than opinions, the estimates of individuals could be statistically averaged. To establish consistency one would hope to find that several persons, using the same information, reached similar inferences independently. Perhaps more specifically consistency could be established by analyses of variance in which certain factors could be controlled. Consistency, based upon common assumptions, evidence, causal

factors and theory, would appear to present a more plausible method for generating future expectations than forcing consensus of opinion."[67]

In this view there is no empirical justification for keeping opinion-makers anonymous (see earlier discussion), just as there is no empirical justification for using dates or probabilities as feedback. Furthermore, there is no logical or empirical justification for seeking convergence under the conditions imposed by Delphi studies.

In modifying Delphi, we should shift entirely away from the idea that convergence improves the accuracy of a forecast. What we need is an instrument to aid the process of clarifying our own assumptions and arguments about the future, as well as those of others. Therefore, there is no reason to use an anonymous questionnaire technique, except perhaps as an evaluative tool to show how changes in estimates reflect the effects of the various arguments presented. Our recent uses of Delphi have been to provide a hypothetical situation to an audience for the purpose of discussing the assumptions that accompany certain claims about the future.

Based on these considerations, a Delphi exercise was developed and conducted at the International Adult Education Seminar, Syracuse University, in December 1969.[68] Participants from the several countries represented at the Conference were first asked to judge when they thought certain events might occur. The events were hypothetical statements constructed from research in progress at the Educational Policy Research Center. Initial judgments of when the events were expected to occur were recorded but not revealed to the group. Small groups were formed and each was asked to discuss, as a working team, several alternative factors that might inhibit or <u>lessen</u> the chances of an event occurring. They were then asked to generate several alternative factors that might <u>increase</u> the likelihood an event would occur.[69] People in each of the groups then discussed the arguments with the entire audience. Finally, a second round

was conducted estimating the expected dates of occurrence for the events discussed. Shifts in dates were evident but no statistical treatment was made. During the course of teaching a graduate seminar,[70] Summer, 1970, additional refinements were made in Delphi.

Below is a listing of specific modifications recommended for the Delphi technique by this writer, based on the various experiences discussed above.

1. Familiarity

   Participants should judge their familiarity with the topics under consideration. (Ament and Gordon used a familiarity scale, at the suggestion of this writer, in a Delphi conducted at IFF, sponsored jointly by EPRC, the State of Connecticut, and others.) Our research shows an effect on the outcome of Delphi studies from familiarity of participants (see Weaver, 1969).

2. C-W Factor

   The assignment of probability factors to the occurrence of events has been dropped altogether for now. Instead, estimates consist simply of an earliest and latest judgment as to when a condition might reach some recognizable proportion. This seems to be the best solution to the confusion between personal confidence and objective probability, as well as a recognition that when people make judgments, they establish in their minds some set of parameters, or what the psychologists call category width (C-W) (see Weaver, 1969).

3. Pedagogy

   Delphi is probably best used as a "conferencing" device where discussion of long-range options might tend to lack focus and organization; therefore, there is no need for the participants to remain anonymous but instead they should be able to confront

each other over issues and assumptions.

4. Focus on Explanation

Participants should be asked to consider and explain why each of several hypothetical conditions might have importance were they to occur; attention is directed but not limited to considerations of magnitude or proportion of the event, and its potential impact on other events.

5. Focus on Underlying Assumptions and Factors

Participants ought to consider at the minimum two sets of factors which might influence the actual occurrence of the events--both a set of negative factors and a set of enhancing factors.

6. Desirability

Participants should weigh the desirability of the events in question, and explain whether their views are connected to critical human issues and values, personal considerations, etc.

7. Feedback

Feedback should consist of the assumptions and arguments generated in (3), (4), (5), and (6) above; the format is open discussion within small groups followed by discussion among groups. (It should be noted that Berghofer used feedback of this sort in his study after consulting an earlier EPRC report.)

8. Convergence

Convergence or divergence which occurs after feedback ought to be taken as an indicator of the force of arguments and clarifying of points of view; in this view it is assumed to have nothing to do with the accuracy of events in question.

The modifications specified above have been incorporated in seminars conducted at the National Educational Technology Conference, New York City, March 1971;[71] Futures Training Project for the State Department of Education, State of Vermont, Spring 1971;[72] New York State System Redesign Project, Cassadaga Valley Schools, Spring 1971.[73]

Finally, it should be clear to the reader that when we speak of Delphi as a pedagogical tool, we do not mean Delphi without substantial changes. In its essentially pure form, Delphi has the same weaknesses as a teaching tool that it suffers as a forecasting tool, namely a lack of explanatory power.

NOTES

1. For a detailed description of the original use of Delphi in technological forecasting see O. Helmer, Social Technology. New York: Basic Books, 1966. For a discussion of the original assumptions underlying the epistemology of Delphi see O. Helmer and N. Rischer. "On the Epistemology of the Inexact Sciences." Management Science, 1959, 6, 25-52.

2. See for instance T. J. Gordon and R. H. Ament, Forecasts of Some Technological and Scientific Developments and their Social Consequences. Middletown, Conn.: Institute for the Future, R-6, September, 1969.

3. Cf. O. Helmer, "The Use of the Delphi Technique in Problems of Educational Innovations," Rand Corporation, P-3499, December 1966. F. R. Cyphert and W. L. Gant, "The Delphi Technique: A Tool for Collecting Opinions in Teacher Education," paper presented at the AERA symposium on Exploring the Potential of the Delphi Technique by Analyzing Its Application, Minneapolis, Minn., March 4, 1970. D. P. Anderson, "Clarifying and Setting Objectives on an Intermediate School District's Objectives Utilizing the Delphi Technique," paper presented at the AERA symposium, op. cit. and N. P. Uhl "Encouraging Convergence of Opinion Through the Use of the Delphi Technique in the Process of Identifying an Institution's Goals," Princeton, New Jersey: Education Testing Service, February 1971.

4. While goals may be perceived by some to be important only because they are personally relevant, the same goals may be held by others to be important on moral grounds. It seems crucial to know which. For instance, A may not rank a goal, say, achieving racial balance in the schools as important, because it does not affect him personally. B may rank the goal high, not because it affects him personally, but because it is important on some moral consideration. C on the other hand, may feel the goal is relatively unimportant, neither on moral grounds nor because of personal relevancy; instead he may feel it simply does not affect a large enough proportion of society.

5. For a discussion of these tools, see S. Sandow. "The Pedagogical Structure of Methods for Thinking about the Future: The Citizen's Function in Planning." Syracuse, New York: Educational Policy Research Center, Working Draft, August 1970.

6. This distinction is somewhat different than that made in R. V. Ayres, Technological Forecasting and Long-Range Planning. New York: McGraw-Hill, 1969. E. Jantsch. Technological Forecasting in Perspective. Paris: OECD, 1966 and J. R. Bright (Ed.) Technological Forecasting for Industry and Government. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968. Particularly, Jantsch makes the distinction between exploratory and normative forecasting techniques, also used in this report, and between intuitive and feedback techniques not used. The weakness in Jantsch's distinctions is that a single technique, Delphi for instance, may be classed in all four of his categories.

7. For research conducted on extrapolation and modeling techniques at the Educational Policy Research Center, Syracuse, see J. A. Henning and A. D. Tussing. "The U.S. Economy Through 2000: Forecasts of Major Macroeconomic Variables," EPRC Working Draft, April 1971, A. D. Tussing and J. A. Henning, "Long-Run Growth of Non-Defense Government Expenditures in the United States," EPRC, August 1970, J. A. Henning and A. D. Tussing. "Income Elasticity of the Demand for Public Expenditures in the United States," EPRC, June 1970, John Henning, "Econometric Models and Forecasting Simplified." Notes on the Future of Education, I, 3 (Summer 1970).

8. R. C. Judd. "Delphi Decision Methods in Higher Education Administration." Manuscript. University of Toledo, October 1971.

9. O. Helmer, op. cit. 1966.

10. Ibid., 1966, p. 22.

11. F. R. Cyphert and W. L. Gant, op. cit.

12. D. P. Anderson, op. cit.

13. N. P. Uhl, op. cit.

14. N. C. Dalkey and D. L. Rourke, "Experimental Assessment of Delphi Procedures with Group Value Judgments." Santa Monica, California: Rand Corporation, R612-ARPA, February 1971.

15. T. J. Gordon and R. Sahr, "Report on Forecasts of Educational Administration" (mimeographed). Middletown, Conn.: Institute for the Future, September 1968.

16. D. E. Berghofer, "General Education in Post-Secondary Non-University Institutions in Alberta." Research Studies in Post-Secondary Education, No. 9, Alberta Colleges Commission, Admonton, Alberta, April 1970.

17. S.C.T. Clarke and H. T. Coutts, "The Future of Teacher Education." Manuscript. Edmonton, Alberta: The University of Alberta, Faculty of Education, manuscript, 1971.

18. D. Hudspeth, "A Long-Range Planning Tool for Education: The Focus Delphi." State University of New York, June 1970.

19. F. Doyle and D. Goodwill, "An Exploration of the Future in Educational Technology." Bell of Canada, January 1971.

20. F. de Brigard and O. Helmer, Some Potential Societal Developments--1970-1990. Middletown, Connecticut: Institute for the Future, R-7, April 1970.

21. The complete series is as follows: R. de Brigard and O. Helmer, op. cit., April 1970; O. Helmer, T. Gordon, S. Enzer, R. de Brigard and R. Rochberg. Development of Long-Range Forecasting Methods for Connecticut: A Summary. IFF Report R-5, Middletown, Connecticut: Institute for the Future (September 1969); T. Gordon and R. Ament, Forecasts of Some Technological and Scientific Developments and Their Societal Consequences, IFF Report R-6, Middletown, Connecticut: Institute for the Future (September 1969); S. Enzer and R. de Brigard. Issues and Opportunities in the State of Connecticut: 1970-2000, IFF Report R-8, Middletown, Connecticut: Institute for the Future (September 1969); S. Enzer, T. Gordon, R. Rochberg and R. Buchele. A Simulation Game for the Study of State Policies, IFF Report R-9, Middletown, Connecticut: Institute for the Future (September 1969); R. Rochberg, T. Gordon, and O. Helmer. The Use of Cross-Impact Matrices for Forecasting and Planning, IFF Report R-10, Middletown, Connecticut, Institute for the Future (September 1969).

22. H. J. Dyck with the assistance of G. J. Emery, "Social Futures for Alberta." Edmonton, Alberta: Westrede Institute, Commissioned by the Human Resources Research Council and the Commission on Educational Planning, 1970.

23. S. Sandow, op. cit., 1970.

24. For an initial description of the Cornell University project, see J. L. Pfeffer, "Preliminary Draft Essays and Discussion Papers on a Conceptual Approach to Designing Simulation Gaming Exercises." Technical Memorandum No. 1 (preliminary draft). Syracuse, New York: Educational Policy Research Center, Syracuse University Research Corporation, October 1968.

25. Umpleby, The Delphi Exploration. A computer-based system for obtaining subjective judgments on alternative futures, Social Implications of Science and Technology Report F-1. Urbana, Illinois: University of Illinois, August 1969.

26. For recent critique of the Cross-Impact Matrix at the Educational Policy Research Center, Syracuse, see M. Folk, "A Critical Look at the Cross Impact Matrix Method," EPRC, RR-5, August 1971. See also T. J. Gordon and H. Hayward, "Initial Experiments with the Cross-Impact Matrix Method of Forecasting," Futures, Vol. 1, No. 2, December 1968; and R. Rochberg, T. J. Gordon, and O. Helmer, "The Use of Cross-Impact Matrices for Forecasting and Planning," Institute for the Future, Riverview Center, Middletown, Connecticut, R-10, April 1970.

27. Cf. R. S. Crutchfield, "Personal and Situational Factors in Conformity to Group Pressure." Paper presented at XV International Congress of Psychology, Brussels, Belgium, August, 1957. Asch, S. E., "Studies of Independence and Conformity: A Minority of One Against a Unanimous Majority," Psychological Monographs, 1956, 70, No. 9 (Whole No. 416).

28. See H. J. Dyck, op. cit.

29. The nature of the questions put to the subjects in the two traditions is significant. The content of the questions reflects the difference in research goals. The psychologists have used mainly personal events in their research designs, whereas, the Operations Research people have been concerned with abstract events. For instance, among the most frequently used techniques in the psychological studies have been Wallace's story-telling and personal future events measures. In the forecasting research events have been entirely of a "non-personal" nature, e.g.. national elections, armament, technologies such as computers and lasers.

30. D. McGregor, "The Major Determinants of the Prediction of Social Events," Journal of Abnormal Psychology, 1938, 33, 179-204.

31. R. M. Campbell, "A Methodological Study of the Utilization of Experts." Unpublished doctoral dissertation, UCLA, 1966.

32. W. T. Weaver, "Delphi: An Empirical Test of Certain Assumptions." Paper presented at the AERA symposium: "Educational Policy Formation and the Future," Minneapolis, Minnesota, March 3, 1970; W. T. Weaver, "An Exploration into the Relationship Between Conceptual Level and Forecasting Future Events." Unpublished doctoral dissertation, Syracuse University, 1969.

33. J. Waldron, "An Investigation into the Relationship among Conceptual Level, Time Delay of Information Feedback and Performance in the Delphi Process. Unpublished doctoral dissertation, Syracuse University, 1970.

34. N. Israeli, "Some Aspects of the Social Psychology of Futurism," The Journal of Abnormal and Social Psychology, 1930, 25, 121-132.

35. Kastenbaum summarized the shortcomings in Israeli's work as (a) failure to present data regarding age, sex, intelligence or education; (b) unenlightening use of percentage statistics which raise more questions than they answer; (c) employment of categories which overlap each other, leave many responses unclassified, and are not related to one another either logically or statistically (of particular concern is the interchangeable use of personalized life goals and non-personal events); (e) clinical naivete in making use of statements provided by the mental hospital and interpretation of patient protocols. See R. J. Kastenbaum, A Preliminary Study of the Dimensions of Future Time Perspective. Doctoral dissertation, University of Southern California. Ann Arbor, Michigan: University Microfilms, 1960, No. 60-394.

36. N. Israeli, "Attitudes to the Decline of the West," Journal of Social Psychology, 1933, 4, 92-101.

37. McGregor, op. cit.

38. Kaplan, op. cit.

39. Helmer, 1966, op. cit.

40. N. Israeli, "The Psychopathology of Time," Psychological Review, 1932, 39, 486-491.

41. N. Israeli, "Wishes Concerning Improbable Future Events. A Study of Reactions to the Future," Journal of Applied Psychology, 1932, 16, 584-588.

42. N. Israeli, "Group Estimates of the Divorce Rate for the Years 1935-1975," Journal of Social Psychology, 1933, 4, 207-222.

43. A. Roberts, "Dogmatism and Time Perspective: Attitudes Concerning Certainty and Prediction of the Future." Unpublished doctoral dissertation, University of Denver, 1959. A. Roberts and R. S. Herrmann, "Dogmatism, Time Perspective, and Anomie," Journal of Individual Psychology, 1960, 16, 67-72.

44. R. J. Bonier, A Study of the Relationship Between Time Perspective and Open-Closed Belief Systems." M.A. thesis, Michigan State University Library, 1957.

45. McGregor, op. cit.

46. H. Cantril, "The Prediction of Social Events," Journal of Abnormal Psychology, 1938, 33, 364-389.

47. Kaplan, op. cit.

48. N. Dalkey, Experiments in Group Prediction. RAND Corporation, P-3820, March 1968. N. Dalkey, Predicting the Future. RAND Corporation, P-3948, October 1968. N. Dalkey, The Delphi Method: An Experimental Study of Group Opinion. Memorandum RM 5888PR. Santa Monica, California: RAND Corporation, June 1969. N. Dalkey, The Delphi Method, II: Structure of Experiments. Memorandum RM 5957-PR. Santa Monica, California: RAND Corporation, June 1969. N. Dalkey, "An Experimental Study of Group Opinion," Futures, 1969, 1, 408-427. N. Dalkey and O. Helmer, "An Experimental Application of the Delphi Method to the Use of Experts," Management Science, 1963, 9, 458-467.

49. Dalkey, 1968, op. cit.

50. B. Brown and O. Helmer, "Improvement in the Reliability of a Consensus Through the Use of Self-rating," in O. Helmer, Social Technology. New York: Basic Books, Inc., 1966.

51. Campbell, op. cit.

52. Weaver, 1969, op. cit.

53. Waldron, 1970, op. cit.

54. Dalkey, see note #48.

55. Brown, op. cit.

56. Cantril, op. cit.

57. Kaplan, op. cit.

58. McGregor, op. cit.

59. Weaver, op. cit.

60. Waldron, op. cit.

61. Campbell, op. cit.

62. Cf. G. M. Vaughan, "The Transituational Aspect of Conforming Behavior," Journal of Personality, 1964, 32, 335-354.

63. The research showed that highly abstract thinkers were considerably more prone to interference from desirability in making estimates in a Delphi experiment than were concrete persons. See Weaver, 1969.

64. For a detailed discussion and summary of the research on future time perspective and forecasting, see W. T. Weaver, "Future Cognition: Contiguities in Human Thought and the Future," Educational Policy Research Center, Syracuse University Research Corporation, working draft, April 1970.

65. By "self-structure" I mean the construct used by O. J. Harvey, D. Hunt, and H. Schroder in Conceptual Systems and Personality Organization. New York: John Wiley and Sons, 1961; O. J. Harvey, Conceptual Systems and Attitude Change," in M. Sherif and C. Sherif, eds., Attitude, Ego Involvement, and Change. New York: John Wiley and Sons, 1969, defined conceptual system as the totality of one's "ties to and definitions of the world." A conceptual system is one's mental structure comprised of: (a) one's differentiated and interrelated concepts of the environment, and (b) one's rules and principles for defining concepts. Concepts and rules function to accomplish a range of goals and help the individual adapt to the environment. Conceptual system is defined therefore as a cognitive structure which has the potential to evolve into a more and more complex set of "rules" and interrelated concepts. This construct has been shown to be remarkable relevant to studies of future perspective. See, specifically, W. T. Weaver, "An Investigation into the Relationship Between Conceptual Level and Forecasting Future Events," doctoral dissertation, Syracuse University, 1969; and J. S. Waldron, "An Investigation into the Relationship Among Conceptual Level, Time Delay of Information Feedback, and Performance in the Delphi Process," unpublished doctoral dissertation, Syracuse University, 1970.

66. See Michael Folk, op. cit., for extensive and critical review of CIM.

67. See Weaver, 1969, op. cit.

68. For notes on the conference, including W. T. Weaver, "Delphi as a Method for Studying the Future: Testing Some Underlying Assumptions," see Essays on the Future of Continuing Education Worldwide. Edited by W. L. Ziegler. Syracuse University Publications in Continuing Education, July 1970.

69. Several questions had been prepared in advance to help the audience develop sets of factors. From a pilot study we found both inhibiting and enhancing factors fell into categories having to do with financial resources, technology, values and attitudes, and organizational constraints.

70. Graduate credit seminar offered by W. T. Weaver at the University of New Hampshire, Department of Education: Long-Range Planning in American Education.

71. The National Educational Technology Conference, Americana Hotel, New York City, March 1971. Sponsored by Educational Technology journal, edited and published by Mr. Lawrence W. Lipsitz.

72. Training Seminars in Long-Range Planning, Spring 1971. Delphi, together with other forecasting tools, was used in two seminars for State of Vermont planning personnel. The overall training program was organized by Robert F. Bundy. Delphi and other tools were introduced by W. T. Weaver.

73. This seminar was conducted by Robert F. Bundy, Educational Planning Consultant, DeWitt, New York. Modifications in Delphi included those mentioned and additional refinements made by Dr. Bundy.

DELPHI STUDIES IN EDUCATION: SELECTED REFERENCES

Anderson, D. P. "Clarifying and Setting Objectives on an Intermediate School District's Objectives Utilizing the Delphi Technique." Paper presented at the AERA Symposium: "Exploring the Potential of the Delphi Technique by Analyzing Its Application." Minneapolis, Minnesota, March 4, 1970.

Berghoffer, D. "General Education in Post-Secondary Non-University Educational Institutions in Alberta." Research Studies in Post Secondary Education, No. 9, Alberta Colleges Commission. Edmonton, Alberta, April 1970.

Clarke, S. C. T., and H. T. Coutts, "The Future of Teacher Education." Edmonton, Alberta: The University of Alberta, Faculty of Education, 1971.

Cochran, W., M. M. Crumley, and B. N. Overly. A Study of Values Employing Delphi Techniques. Commerce, Texas: East Texas State University, 1970.

Cyphert, F. R. and W. L. Gant. "The Delphi Technique: A Case Study," Phi Delta Kappan. Vol. 52, No. 5, January 1971, 272-234.

Cyphert, F. R. and W. L. Gant. "The Delphi Technique: A Tool for Collecting Opinions in Teacher Education," Journal of Teacher Education. Vol. 21, No. 3, Fall, 1970, pp. 419-420.

Cyphert, F. R. and W. L. Gant. "The Delphi Technique: A Tool for Collecting Opinions in Teacher Education." Paper presented at the AERA Symposium: "Exploring the Potential of the Delphi Technique by Analyzing Its Application." Minneapolis, Minnesota, March 4, 1970.

Dalkey, N. and D. L. Rourke. Experimental Assessment of Delphi Procedures With Group Value Judgments. Santa Monica, California: RAND Corporation R-612-ARPA, February 1971.

de Brigard, Raul and Olaf Helmer, Some Potential Societal Developments: 1970-2000, Report R-7. Institute for the Future, April 1970.

Doyle, F. and D. Goodwill. An Exploration of the Future in Educational Technology. Bell of Canada, January 1971.

Dyck, H. J., with the assistance of G. J. Emery. "Social Failures. Albert" Edmonton, Canada: Westrede Institute, 1970.

Goodey, Brian. "Perception, Gaming and Delphi." Paper for Architectual Psychology Conference, Kingston Polytechnic, Surry, England, September 3, 1970.

Gordon, T. J. and R. Sahr. "Report on Forecasts of Educational Administration" (mimeographed preliminary draft). Middletown, Connecticut: Institute for the Future. September 1969.

Helmer, O. "The Use of the Delphi Technique in Problems of Educational Innovations." P-3499. Santa Monica, California: RAND Corporation, December 1966.

Hudspeth, D. "A Long-Range Planning Tool for Education: The Focus Delphi." State University of New York, June 1970.

Jacobson, James A. "Forecasting Future Developments in Education." A paper presented at the American Education Research Association Meeting: Minneapolis, Minnesota, March 4, 1970.

Judd, Robert C. "Delphi Decision Methods in Higher Education Administration." Manuscript. University of Toledo, October 1971.

Judd, R. C. "Delphi Method: Computerized 'Oracle' Accelerates Consensus Formation." College and University Business. Vol. 49, No. 3, September 1970, pp. 30-34.

Matsuskita, Hiroshi. "Diffusion of Soft Technology." Prepared by Moboru Makino for Education for Future: International Future Research Conference, Kyota, Japan, 1970.

Sandow, Stuart A. "The Pedagogy of Planning: Defining Sufficient Futures." Futures: Journal of Forecasting and Planning, December 1971, III, 4.

Sandow, Stuart A. Emerging Policy Issues in Law. Syracuse, New York: Educational Policy Research Center, 1971.

Turoff, Murray. "Delphi and Its Potential Impact on Information Systems." Prepared for the Fall Joint Computer Conference, November 1971.

Turoff, Murray. "Delphi Conferencing," TM-125. Executive Office of the President, Office of Emergency Preparedness, Office of the Assistant Director for Research Analysis. Washington, D. C., March 1971.

Uhl, Norman P. Encouraging Convergence of Opinion Through the Use of the Delphi Technique in the Process of Identifying an Institution's Goals. Princeton, New Jersey: Education Testing Service, February 1971.

Umpleby, S. The Delphi Exploration. A computer based system for obtaining subjective judgments on alternative futures. Social Implications of Science and Technology Report F-1. Urbana, Illinois: University of Illinois, August 1969.

Waldron, J. "An Investigation into the Relationship Among Conceptual Level, Time Delay of Information Feedback and Performance in the Delphi Process." Unpublished doctoral dissertation. Syracuse University, 1970.

Weaver, W. T. "An Exploration into the Relationship Between Conceptual Level and Forecasting Future Events." Unpublished doctoral dissertation. Syracuse University, 1969.

Weaver, W. T. "Delphi: An Empirical Test of Certain Assumptions." Paper prepared at the AERA Symposium: Educational Policy Formation and the Future. Minneapolis, Minnesota, March 3, 1970.

Weaver, W. T. "Delphi as a Method for Studying the Future: Testing Some Underlying Assumptions." Paper prepared for the International Seminar on Adult Education. Syracuse University, December 1969.

Weaver, W. T. "The Delphi Forecasting Method," Phi Delta Kappan. January 1971, pp. 267-272.

Welty, Gordon. "A Critique of Some Long-Range Forecasting Developments." Proceedings of the 38th Session of the International Statistical Institute: Washington, D. C. August 10-20, 1971, pp. 404-409.

Uhl, Norman. "Encouraging Convergence of Opinion Through the Use of Delphi Technique, in the Process of Identifying an Institutions Goals." Paper presented at the American Educational Research Association, New York City, February 4-7, 1971.

SELECTED RELATED STUDIES

Ament, Robert H., "Comparison of Delphi Forecasting Studies in 1964 and 1969," Futures, Vol. 2, No. 1, March 1970.

Brown, B., S. W. Cochran, and N. C. Dalkey, The Delphi Method II: Structure of Experiments. RM-5957-PR. The RAND Corporation, June 1969.

Campbell, Robert M. "A Methodological Study of the Utilization of Experts in Business Forecasting." Unpublished doctoral dissertation. University of California, Los Angeles, September 1966.

Dalkey, N. The Delphi Method II: Structure of Experiments. Memorandum RM 5957-PR. Santa Monica, California: RAND Corporation, June 1969.

Dalkey, N. C. "Analyses From a Group Opinion Study," Futures, Vol. 1, No. 6, December 1969, pp. 541-551.

Dalkey, N. C., and Olaf Helmer. An Experimental Application of the Delphi Method to the Use of Experts. RM-727-PR (Abridged), The RAND Corporation, July 1962. Also published in Management Science, Vol. 9, No. 3, April 1963, pp. 458-467.

Dalkey, N. C. "An Experimental Study of Group Opinion: The Delphi Method," Futures, Vol. 1, No. 5, September 1969, pp. 408-426.

Dalkey, N. The Delphi Method: An Experimental Study of Group Opinion. Memorandum RM 5888-PR. Santa Monica, California: The RAND Corporation, June 1969.

Dalkey, N. C. Delphi. P-3704. The RAND Corporation, October 1967.

Dalkey, N. Experiments in Group Prediction. P-3820. The RAND Corporation, March 1968.

Dalkey, N. Predicting the Future. P-3948. The RAND Corporation, October 1968.

Dalkey, N. C., B. Brown, and S. W. Cochran. The Delphi Method III: Use of Self-Ratings to Improve Group Estimates. RM-6115 PR, The RAND Corporation, October 1969.

Dalkey, N. C., B. Brown, and S. W. Cochran. The Delphi Method IV: Effect of Percentile Feedback and Feed-In of Relevant Facts, RM-6118-PR. The RAND Corporation, March 1970.

Enzer, Slewyn. A Case Study Using Forecasting as a Decisionmaking Aid. Working Paper WP-2. Institute for the Future, December 1969.

Enzer, Selwyn, and Raul de Brigard. Issues and Opportunities in the State of Connecticut: 1970-2000. Report R-8. Institute for the Future, March 1970.

Gordon, Theodore J. and Robert H. Ament. Forecasts of Some Technological and Scientific Developments and Their Societal Consequences. Report R-6. Insitute for the Future, September 1969.

Helmer, Olaf. Analysis of the Future: The Delphi Method. P-3558. The RAND Corporation, March 1967.

Helmer, Olaf. Convergence of Expert Consensus Through Feedback. P-2973. The RAND Corporation, September 1964.

Helmer, Olaf, Theodore J. Gordon, Selwyn Enzer, Raul de Brigard, and Richard Rochberg. Development of Long-Range Forecasting Methods for Connecticut: A Summary. Report R-5. Institute for the Future, September 1969.

Helmer, Olaf and Bernice Brown. Improving the Reliability of Estimates Obtained from a Consensus of Experts. Appendix to Olaf Helmer, Social Technology, New York: Basic Books, 1966.

Helmer, Olaf. Long-Range Forecasting--Roles and Methods. Paper P-7. Institute for the Future, May 1970.

Helmer, Olaf. Methodology of Societal Studies. P-3611. The RAND Corporation, June 1967.

Helmer, Olaf. New Problems in Early Forecasting of Public Problems. P-3576. The RAND Corporation, April 1967.

Helmer, Olaf and Nicholas Rescher. On the Epistemology of the Inexact Sciences. P-1513. The RAND Corporation, 1959. Also published in Management Science, Vol. 6, 1959.

Helmer, Olaf. Social Technology. New York: Basic Books, 1966.

Helmer, Olaf. The Future of Science. P-3607. The RAND Corporation, May 1967.

Helmer, Olaf. Systematic Use of Expert Opinions. P-3721. The RAND Corporation, November 1967.

Jantsch, E. Technological Forecasting in Perspective. Paris: Organisation for Economic Co-operation and Development, 1967.

Kaplan, A., A. Skogstad, and M. Girschick. "The Prediction of Social and Technological Events," Public Opinion Quarterly, 1950, pp. 14, 93-110.

North, Harper Q. and Donald L. Pyke. "'Probes' of the Technological Future," Harvard Business Review, May-June 1969.

Pill, Juri. "The Delphi Method: Substance, Content, A Critique and Annotated Bibliography." Operations research Department, Case Western Reserve University, Technical Memorandum #183, 1970.

Quade, E. S. Cost Effectiveness: Some Trends in Analysis. P-3529. The RAND Corporation, March 1967.

Sahr, Robert. A Collation of Similar Delphi Forecasts. Working Paper WP-5. Institute for the Future, April 1970.

Sulc, Oto. Forecasting the Interactions Between Technological and Social Changes. Report 061-237-4571. Manchester Business School, University of Manchester, July 1968.